

Presidenza del Consiglio dei Ministri



**INTELLIGENZA ARTIFICIALE E MEDICINA:
ASPETTI ETICI**

29 maggio 2020

INDICE

Presentazione	3
1. Definizioni di IA e recenti sviluppi	5
2. Sviluppi della IA nell'ambito della salute.....	7
3. Problemi etici emergenti	9
3.1. IA nella relazione paziente-medico.....	9
3.2. L'affidabilità della IA e la opacità degli algoritmi.....	10
3.3. IA e dati: tra privacy e condivisione dei dati	13
3.4. Consenso e autonomia.....	14
3.5. La responsabilità.....	14
3.6 La formazione in ambito medico, tecnologico e sociale.....	15
4. Raccomandazioni	16

Presentazione

Nell'ambito di un incontro del Presidente del Consiglio Conte al CNB (26 settembre 2019) è stata avanzata dal Presidente una specifica richiesta di un pronunciamento del Comitato sull'utilizzo della Intelligenza artificiale (IA) nel campo della salute. Nell'ambito del discorso rivolto al CNB, il Presidente Conte ha sottolineato: "Nella prospettiva cui ci stiamo avviando, dove l'innovazione tecnologica caratterizzerà ancor di più il nostro vivere quotidiano e dove ovviamente, in una prospettiva di governo, premeremo ancor di più per avere una svolta in senso digitale, è chiaro che la intelligenza artificiale e la robotica giocheranno un ruolo ancora maggiore e ci sfiderà a confrontarci con i dilemmi morali (...). Per cui vi chiederei su questo di accompagnarmi con una riflessione nel momento stesso in cui ci avviamo in quella direzione di sviluppo".

Al fine di elaborare una risposta al quesito con riferimento alle applicazioni nell'ambito della salute e della medicina il CNB ha costituito un gruppo misto con il CNBBSV, coordinato dai Proff.: Salvatore Amato, Carlo Casonato, Amedeo Cesta, Roberto Cingolani, Lorenzo d'Avack, Silvio Garattini, Laura Palazzani.

Il documento è stato curato dalla Prof. Laura Palazzani.

Il parere, a partire da una definizione di IA, ne analizza le origini e gli sviluppi più recenti, con specifico riferimento alla enorme disponibilità di dati e alla potenza di calcolo. Il documento sottolinea le opportunità e i rischi della IA e le principali applicazioni in medicina, anche nell'ambito della pandemia Covid-19.

I Comitati, nel contesto della rapida evoluzione di queste tecnologie, richiamano alcuni elementi di riflessione critica per una comprensione e valutazione della IA. Nell'ambito della relazione paziente-medico, sottolineano, da un lato, le opportunità che possono consentire agli operatori sanitari di ridurre il tempo per attività burocratiche, di routine o pericolose, permettendo loro di avere più disponibilità nella relazione di cura per il paziente, dall'altro lato, descrivono i rischi che una "assistenza cognitiva automatizzata" possa ridurre le abilità del medico e dell'operatore sanitario. Il documento sottolinea la rilevanza di strumenti che garantiscano l'affidabilità della IA, mediante validazione, e riducano, nei limiti del possibile, la opacità, gli errori e le possibili discriminazioni dovute a cause tecnologiche e/o umane. Dato l'enorme uso di dati, è indispensabile inoltre una adeguata tutela della privacy, anche considerando la possibilità della condivisione dei dati come "bene sociale".

Il consenso informato, seppur con alcune difficoltà, date da una informativa del medico e da una comprensione del paziente non sempre semplici e abituali, rimane un elemento essenziale della relazione paziente-medico. Particolare attenzione è pertanto dedicata anche ad una nuova formazione in ambito medico, tecnologico e sociale. In tal senso, i Comitati ritengono indispensabile ripensare la formazione dei professionisti della salute in modo dinamico, con una revisione flessibile dei programmi di studio da parte di commissioni interdisciplinari, per combinare in modo trasversale ed interdisciplinare le diverse competenze di IA e, al tempo stesso, introdurre la rilevanza dell'etica nei corsi di formazione degli ingegneri, informatici, sviluppatori, con particolare riferimento all'etica nella progettazione e applicazione delle tecnologie. Obiettivo importante dovrebbe essere anche la promozione di una coscienza pubblica della società su opportunità e rischi delle nuove tecnologie, oltre che un aggiornamento normativo sui profili concernenti la responsabilità nell'applicazione delle nuove tecnologie e la promozione della ricerca su IA in ambito sia privato, sia pubblico.

L'identificazione della responsabilità, sul piano giuridico, esige una verifica delle categorie esistenti, dato il pluralismo delle competenze tra il progettista, il venditore del software, il proprietario, l'utilizzatore (il medico) o terze parti.

Sono state organizzate alcune audizioni di esperti interni ai due Comitati: il Dott. Amedeo Cesta, membro del CNB (13 dicembre 2019) e il Prof. Roberto Cingolani, membro del CNBBSV (31 gennaio 2020) ed è stato invitato un esperto esterno il dott. Alberto Tozzi, Responsabile Unità di funzione medicina digitale e telemedicina, Responsabile Area di Ricerca malattie multifattoriali e malattie complesse, Ospedale Pediatrico Bambino Gesù (29 Maggio 2020); il Prof. Carlos Romeo Casabona, Professore di diritto costituzionale presso Università dei Paesi Baschi, Rappresentante della Spagna presso Il Comitato di Bioetica del Consiglio d'Europa, Membro dell'European Group on Ethics in Science and New Technologies presso la Commissione Europea, Visiting Professor Università di Roma Tor Vergata (13 dicembre 2019).

Il documento è stato votato dal CNB nella plenaria del 29 maggio 2020 all'unanimità dei presenti: Proff. Salvatore Amato, Luisella Battaglia, Carlo Caltagirone, Stefano Canestrari, Cinzia Caporale, Carlo Casonato, Francesco D'Agostino, Bruno Dallapiccola, Antonio Da Re, Lorenzo d'Avack, Mario De Curtis, Gianpaolo Donzelli, Silvio Garattini, Mariapia Garavaglia, Marianna Gensabella, Assunta Morresi, Laura Palazzani, Tamar Pitch, Lucio Romano, Massimo Sargiacomo, Monica Toraldo di Francia e Grazia Zuffa.

Il Dott. Maurizio Benato, componente delegato della FNOMCeO, la Dott.ssa Carla Bernasconi, componente delegato della FNOVI, il Dott. Amedeo Cesta, componente delegato del Presidente del CNR, la Dott.ssa Paola Di Giulio, componente delegato del Presidente del CSS, il Prof. Carlo Petrini, componente delegato del Presidente dell'ISS, pur non avendo diritto al voto, hanno voluto tuttavia esprimere la loro adesione al parere.

Assenti alla seduta hanno successivamente espresso la loro adesione al parere i Proff.: Riccardo Di Segni, Luca Savarino e Lucetta Scaraffia.

È stato altresì votato dal CNBBSV, in data 5 giugno 2020, dai Proff.: Andrea Lenzi (Presidente), Antonio Amoroso, Antonio Bergamaschi, Carlo Caltagirone, Roberto Cingolani, Fabio Fava, Paolo Gasparini, Maurizio Genuardi, Marco Gobbetti, Paola Grammatico, Mauro Magnani, Piero Angelo Morandini, Luigi Naldini, Ferdinando Nicoletti, Giuseppe Novelli, Pierfranco Pignatti, Roberta Siliquini, Paolo Visca.

1. Definizione di IA e recenti sviluppi

Non è scontato dare una definizione omogenea di Intelligenza Artificiale (IA) soprattutto alla luce dei recenti sviluppi che hanno portato ad un uso molto diffuso del termine¹. In generale, con tale termine può identificarsi il settore delle Tecnologie della Informazione e della Comunicazione (ICT, Information and Communication Technologies) che ha l'obiettivo di imitare con tecnologie informatiche alcuni aspetti dell'intelligenza umana, per sviluppare "prodotti informatici o macchine" in grado sia di interagire e di apprendere dall'ambiente esterno, sia di assumere decisioni con crescenti gradi di autonomia.

Ciononostante, il termine "intelligenza", coerentemente con il termine inglese "intelligence", non designa qualità propriamente umane conferite alle macchine, ma descrive funzioni che rendono alcuni comportamenti delle macchine simili a quelli di un essere umano. L'altro aspetto assolutamente particolare è che ci troviamo, per la prima volta, dinanzi a tecnologie che agiscono sistematicamente come utenti autonomi di altre tecnologie.

Gli sviluppi della IA, dopo un periodo di stasi, sono ripresi progressivamente in questi ultimi decenni² per arrivare all'impatto significativo attuale. Il grande interesse verso l'IA, che ha fatto sì che i risultati, per anni fondamentalmente relegati al campo della ricerca scientifica, siano approdati prima nei laboratori industriali e da qui al mercato, ha delle spiegazioni nel progresso tecnologico.

Per interpretare il momento che stiamo vivendo dobbiamo osservare due aspetti legati all'evoluzione della tecnologia: la potenza di calcolo degli elaboratori, che ha reso possibile che brillanti risultati di ricerca che pochi anni fa risolvevano "istanze giocattolo"³ di un problema, grazie a tale potenza sono arrivati a risolvere istanze di problemi reali. Il secondo aspetto è legato alla accresciuta disponibilità di enormi quantità di dati⁴ e informazioni (che costituiscono gli "esempi" o elementi di base per la macchina), tratte da ICT e dal web, e alla elaborazione di algoritmi⁵ che ora sono diventati "eseguibili": dati e algoritmi costituiscono la capacità di "apprendimento" (*machine learning*), che consente alla macchina, sulla base delle informazioni immagazzinate e archiviate (i dati), di scoprire relazioni nascoste tra i dati e la connessione delle informazioni (gli algoritmi)⁶. Gli algoritmi integrano espressioni matematiche, che trovano applicazione in problemi quotidiani per trovare associazioni, identificare tendenze

¹ Cfr. il Technical report del Joint Research Center della Commissione europea su *Defining Artificial Intelligence. Towards and Operational Definition and Taxonomy of AI* del 27 febbraio 2020 (<https://op.europa.eu/en/publication-detail/-/publication/6cc0f1b6-59dd-11ea-8b81-01aa75ed71a1/language-en/format-PDF/source-118380790>).

² Nella IA classica si ragionava con una matematica legata alla logica filosofica, traendo delle deduzioni sulla base della causalità. Oggi, invece, il problema è che la conoscenza possa essere astratta direttamente dai dati.

³ Fra le prime la partita a scacchi fra il calcolatore *Deep Blue* e l'allora campione del mondo Kasparov.

⁴ Sul tema dei dati cfr. CNB, *Tecnologie dell'informazione e della comunicazione e big data: profili bioetici* (2016).

⁵ L'algoritmo è una successione in sequenza di istruzioni che definiscono le operazioni elementari che la macchina deve eseguire sui dati per ottenere i risultati. È un procedimento sistematico di calcolo che risolve un determinato problema attraverso un numero finito di istruzioni elementari con una quantità di dati finita.

⁶ Vanno distinti, fra l'altro, due tipi di IA: a) IA debole emulazionistica (*weak AI*) basata sul principio che l'essenza del funzionamento del cervello non risiede nella sua struttura ma nelle sue prestazioni; b) IA forte, simulazionistica (*strong AI*) basata sulla ricerca per riprodurre il più fedelmente possibile la fisiologia del cervello, possibile estensione di quella debole.

ed individuare regolarità all'interno di qualunque insieme di dati, alla base dei comportamenti umani, espressi dai dati e dalle informazioni⁷.

Gli algoritmi hanno bisogno di una fase di addestramento a partire da esempi forniti dall'uomo che costituiscono la base da cui imparare⁸. Questi due effetti combinati (dati e algoritmi) sono alla base del progresso attuale: mentre in precedenza i programmi di IA “erano addestrati a mano” dalla abilità dei loro creatori, ora possono anche “addestrarsi da soli”, con dei limiti, e questo ha aperto nuove strade. Nel caso dei dati disponibili su internet, si può programmare l'acquisizione automatica dei dati man mano che diventano disponibili (alcuni strumenti di marketing o di management finanziario usano questo meccanismo).

Sulla base di questo “addestramento” l'IA è in grado di predire, con diversi gradi di probabilità. La correttezza nella predizione della IA è proporzionale al numero e alla qualità e accuratezza dei dati inseriti ed esperienze immagazzinate su un determinato tema, ma potrebbe sempre fallire se dovesse emergere un caso mai capitato prima. Le predizioni della IA dipendono dai dati e dagli algoritmi con cui il sistema viene “addestrato”: le predizioni possono essere “sbagliate” per l'imprecisione dei dati forniti o per l'uso di assunzioni non fondate. È necessario dunque curare il “nutrimento” o “addestramento” dell'IA: in tale ambito l'uomo mantiene un ruolo centrale. In questo senso non si può – ad oggi – parlare di “autonomia decisionale” della macchina.

Questo progresso ha riguardato le aree dell'IA in misura differente: molto i compiti percettivi (la visione artificiale, il riconoscimento di specifici oggetti, l'interpretazione del linguaggio parlato e scritto), in maniera minore i compiti simili al ragionamento umano, quali sviluppare la capacità di ragionare, capire le intenzioni, elaborare argomentazioni, oltre che costruire un discorso verbale molto articolato. È oggetto di ampia discussione la c.d. “capacità logica” della IA: quello che sembra un processo logico-deduttivo basato sulla concatenazione tipica del ragionamento umano (di associazione causale che parte dalla descrizione della realtà e inferisce conclusioni) eseguito dalla IA è, in verità, un modello dinamico, basato sul velocissimo confronto e correlazione con esempi immagazzinati. La IA, in questo senso, applica un principio matematico che evince correlazioni tra i dati, ma non ragiona in modo “logico” in senso proprio. Le correlazioni comunque sono sempre predizioni probabilistiche e contengono molti limiti ed eccezioni.

Una particolare area del *machine learning* è il *deep learning* derivante da un processo di imitazione del cervello umano, basato sulla creazione di reti di neuroni artificiali. Nel *deep learning*, la macchina estrae dei significati ragionando su grosse quantità di dati: si tratta di metodi di apprendimento “automatico” (più che “autonomo”), che, al momento, generano risultati “opachi” non facilmente

⁷ L'IA non è una scoperta scientifico-tecnologica nuova ma un'area di ricerca la cui nascita generalmente si fa risalire al 1956 (anno della Dartmouth Summer Research Project on Artificial Intelligence a cui sicuramente risale il conio di IA), ma che risale ancora prima agli studi sul neurone artificiale di J. McCULLOCH (*A Logical Calculus of the Ideas Immanent in Nervous Activity*, in “Bulletin of Mathematical Biophysics”, 1943, 5 (4), pp. 115–133) e agli studi di Alan Turing (*Computing Machinery and Intelligence*, in “Mind”, 1950). L'area di ricerca ha attraversato vari periodi: un primo momento di auge si è avuto negli anni ottanta in cui si ebbe un primo tentativo di commercializzazione (legato ai così detti “sistemi esperti”) che si rivelò un successo limitato e ben presto deludente. Seguì un lungo periodo di rallentamento nei suoi avanzamenti (identificato anche come “inverno della IA”) fondamentalmente dedicato ad un ritorno nei laboratori per riprendere studi di base, anche fortemente teorici (A. VESPIGNANI, *L'algoritmo e l'oracolo*, Milano 2020).

⁸ Un'equazione nella quale a sinistra dell'uguale viene accumulato un certo numero di esempi.

spiegabili (il problema della black box). Questi portano a risultati superiori in prestazioni, settorialmente avvicinabili a quelli umani, pur con alcuni limiti.

Gran parte dei timori e delle incertezze attuali sull'IA si basano sulla presupposizione di situazioni oggi ancora irrealistiche, come la “sostituzione” della capacità decisionale dell'umano, o la “autonomizzazione” della macchina che può sfuggire al controllo dell'umano⁹. Per quanto teoricamente possibile, siamo ancora lontani da questo scenario. L'IA è uno strumento potente, ma accessorio alla decisione umana. Il problema oggi pertanto non è tanto quello – da più parti sottolineato – dei timori per la “autonomia” della IA, ma semmai il problema riguarda il fatto che un sistema esperto che diventa ottimale nel suggerire “decisioni” all'uomo rischia di ridurre l'attenzione umana con la possibile conseguenza della riduzione delle abilità umane (o deskilling)¹⁰. In questo senso è importante una riflessione sulla sinergia tra uomo e macchina, e sulla ricerca delle modalità di “supporto” intelligente che consenta all'uomo un “controllo umano significativo” in termini di supervisione e attenzione.

2. Sviluppi della IA nell'ambito della salute

2.1. Fin dagli anni '70 la IA è stata considerata come “area emergente” (detta area dei sistemi esperti) capace di svolgere ragionamenti sulla conoscenza limitata per imitare i ragionamenti medici. Oggi nella medicina esistono già molte applicazioni dell'IA finalizzate a migliorare la pratica sanitaria.

Così, la IA può assistere il professionista nella prevenzione e nel classificare e stratificare le condizioni del paziente (riducendo l'incertezza diagnostica¹¹); nel comprendere perché e come i pazienti sviluppano le malattie (riducendo l'incertezza fisiopatologica); nel considerare quale trattamento sarà più appropriato per loro (riducendo l'incertezza terapeutica); nel prevedere se si riprenderanno con o senza un trattamento specifico (riducendo l'incertezza prognostica e incremento della predizione della insorgenza o evoluzione delle patologie), e così via.

⁹ Si tenga conto che queste preoccupazioni di trovano in diverse dichiarazioni internazionali. Ad esempio, “the specific characteristics of many AI technologies, including opacity ('black box-effect'), complexity, unpredictability and partially autonomous behaviour, may make it hard to verify compliance with, and may hamper the effective enforcement of rules of existing EU law meant to protect fundamental right” (European Commission, *White Paper On Artificial Intelligence - A European approach to excellence and trust*, 19.02. 2020, p. 12). Il documento dell'Unesco su *Robotics Ethics* del 14, settembre, 2017 suggerisce di distinguere tra i tradizionali computer deterministici e i computer stocastici o probabilistici. L'*Ethics Guidelines for Trustworthy AI*, elaborato dall'High-Level Expert Group on Artificial Intelligence nell'aprile 2019 sotto l'egida dell'Unione europea suggerisce la necessità di preservare la dimensione “umano-centrica” (*human-centric*) delle nuove tecnologie. Anche l'European Group on Ethics in Science and New Technologies nel documento del 20 marzo 2018 su *Artificial Intelligence, Robotics and 'Autonomous' Systems* sottolinea l'importanza “that humans - and not computers and their algorithms - should ultimately remain in control, and thus be morally responsible”.

¹⁰ È ciò che è accaduto nel trasporto aereo: l'assistenza 'intelligente' al pilota rischia di fare diminuire l'attenzione umana e, per ovviare a questo problema, il pilota viene frequentemente addestrato su un simulatore.

¹¹ Il sistema Dermosafe, attualmente utilizzato in molteplici ospedali, ha buone probabilità di individuare le situazioni critiche ed intervenire tempestivamente per arginare lo sviluppo del tumore. E l'integrazione artificiale “*embodied*” può dare effetti positivi nell'interazione col paziente nella prevenzione delle malattie degenerative.

Oggi, inoltre, si sta cercando di elaborare un supporto al medico e agli operatori sanitari¹² affinché sia possibile consultare le linee guida più aggiornate e appropriate, anche mentre lavorano in corsia¹³. Un corretto impiego della IA potrebbe anche migliorare e rendere più efficienti e scorrevoli i flussi all'interno delle strutture sanitarie, dal triage alla gestione delle emergenze o degli interventi di selezione dei dispositivi medici, sia nel settore farmaceutico in relazione, tra l'altro, al possibile impiego di packaging intelligente al confezionamento ed alla vendita del farmaco. Inoltre, la IA può essere utilmente impiegata per la sperimentazione clinica e nella prospettiva della medicina di precisione.

2.2. Come già avvenuto in riferimento alla diffusione dell'Ebola, anche nella pandemia Covid-19, la IA viene utilizzata per dare un contributo decisivo alla lotta contro il virus Sars-Co V-2, che ha colpito il mondo dalla fine del 2019. Trattandosi di un'epidemia e con una diffusione repentina, su una scala globale, è indispensabile avere strumenti altrettanto veloci e applicabili simultaneamente in aree diverse. Molteplici gli usi fatti dell'IA e tra questi indichiamo come prevalenti: osservazione e predizione dell'evoluzione degli andamenti della pandemia; finalità diagnostiche della patologia; ricerca di un vaccino o di una cura (conoscendo l'attività dei farmaci sulle catene metaboliche e la struttura dei recettori sui quali si vuole agire, l'IA può essere di supporto); assistenza al personale sanitario e ai pazienti, fornendo farmaci e cibo e misurando i segni vitali; disinfezione e decontaminazione; condivisione della conoscenza e rilevazione della disinformazione; controllo e tracciabilità dei comportamenti della popolazione¹⁴.

2.3. In medicina la IA, sulla base dei dati e dell'analisi dei dati, esplica i propri effetti nell'interpretazione di pattern complessi. Ad esempio, nell'interpretazione delle immagini l'IA riconosce segnali che non sono distinguibili dall'occhio umano. Se l'esperienza dell'IA su questi segnali è grande, l'IA avrà una buona performance. Molte le applicazioni nell'ambito della medicina personalizzata, dell'analisi di big data in genomica, nella sperimentazione dei farmaci¹⁵, nel funzionamento di robot chirurgici (realtà virtuale, teleoperazione,

¹² Nel testo quando si fa riferimento al medico, ove rileva, l'indicazione include anche il riferimento agli operatori sanitari.

¹³ Attualmente, IBM sta cercando di realizzare una versione di Watson che sia utile al medico; in particolare, un sistema che fornisca al medico, mentre sta operando, delle specifiche informazioni presenti nella letteratura di quell'ambito. Ogni mese la Food and Drug Administration si trova a valutare numerosi algoritmi di diagnostica per immagini. Sempre nel campo medico la FDA ha proposto varie Linee guida e Protocolli per garantire le *best practice* nella gestione di questi algoritmi che evolvono nel tempo.

¹⁴ Si veda il documento del Consiglio d'Europa *IA e lotta contro il coronavirus Covid-19*: <https://www.coe.int/en/web/artificial-intelligence/ia-e-lotta-contro-il-coronavirus-covid-19>. È particolarmente significativa l'iniziativa delle strutture sanitarie di Bergamo di nuovi macchinari che, con l'aiuto di finanziamenti privati, si sono dotate di robot per automatizzare le procedure che solitamente sono eseguite dagli operatori, per preparare i campioni di saliva e muco prelevati tramite i tamponi prima di essere inseriti negli scanner che rilevano l'eventuale presenza del coronavirus. Questo sistema automatico si rivela utile non solo per alleggerire parte del lavoro dei tecnici di laboratorio, che si possono dedicare ad altre attività, ma anche per raddoppiare il numero di tamponi analizzati giornalmente.

¹⁵ Si pensa che sia possibile accorciare il tempo dedicato alla scoperta di un farmaco da 4, 5 anni a uno, con un taglio dei costi dell'80 %. Ad esempio, Halicina è il primo antibiotico scoperto da un algoritmo di intelligenza artificiale. È un antibiotico ad ampio spettro che agisce su batteri difficili da trattare e resistenti agli antibiotici. Un antitumorale, BPM31510, ottenuto con sistemi di IA che

analisi di immagini in tempo reale). I sistemi robotici sono dei sistemi integrati che fanno riferimento all'umano, integrando capacità computazionali (il pensiero), immagazzinamento di dati (la memoria), apparati sensoriali (la vista, l'udito, il tatto) e gli attuatori per generare azioni fisiche (il sistema muscolo scheletrico). La sinergia di questi fattori viene applicata all'interno del robot per operare secondo le necessità dettate dall'operatore (il medico). La crescita della IA nell'ambito della salute e nei diversi settori è data anche negli ultimi decenni dalla crescita della robotica in campo civile, adoperata non solo per uso domestico e ludico (social robot), ma anche in quello sanitario-medico sia con funzioni diagnostiche che per la pratica clinica in senso stretto. Anche in questo campo le finalità sono molteplici: attività di supporto chirurgico¹⁶; in ambito diagnostico e di prevenzione (con l'ausilio di nanorobots); nell'ambito della cura, della riabilitazione e dell'assistenza personale ad anziani, a persone con difficoltà motorie, a persone affette da autismo¹⁷.

3. Problemi etici emergenti

Il Comitato prende atto dell'evoluzione rapida che sta avvenendo nella medicina e apprezza gli enormi progressi e le straordinarie opportunità dischiuse dalla IA. Il processo è avviato e la trasformazione sembra essere prorompente, inevitabile ed irreversibile.

In questo contesto in evoluzione e in transizione, il Comitato intende richiamare alcuni elementi di riflessione etica, senza né esaltare né ostacolare lo sviluppo della tecnologia, bensì offrire una riflessione "critica" per una comprensione e valutazione delle nuove tecnologie, per cercare di comprendere come veramente "funzionano" e valutare ciò che acquistiamo in termini di potenzialità. L'obiettivo è identificare le condizioni etiche per uno sviluppo della IA che non rinunci ad alcuni aspetti della nostra umanità, in un nuovo "umanesimo digitale", per una medicina "con" le macchine e non "delle" macchine. Nella consapevolezza che è l'uomo che costruisce la tecnologia e che la tecnologia non è uno strumento neutrale, in quanto modifica inevitabilmente la stessa relazione paziente-medico.

3.1. IA nella relazione paziente-medico

L'impatto sui principi bioetici dipende tanto dai settori di applicazione dell'IA nel campo sanitario quanto dall'identità del medico alla luce di questi nuovi supporti e dalla funzione complessiva che si affida al servizio sanitario.

L'uso di macchine intelligenti e robot in medicina, nella misura in cui sono e saranno più efficienti, precisi, rapidi e meno costosi, sembra auspicabile se si considera tale sostituzione dell'uomo con riferimento ad attività ripetitive, noiose, pericolose, umilianti o faticose. Se correttamente utilizzata, la IA potrebbe ridurre i tempi che i professionisti devono dedicare ad attività burocratiche, di mera

setacciano migliaia di tessuti umani, ha superato un trial di fase 2 per pazienti con tumore pancreatico in stato avanzato.

¹⁶ Il robot-surgeon, dotato di forme di IA, svolge una innovativa attività di supporto, anche permettendo di eseguire l'operazione chirurgica a mezzo di "realtà virtuale", mediante l'ausilio di una penna collegata ad un robot piccolissimo che indirizza il fascio laser nella direzione voluta dal chirurgo. Pertanto, il medico si trova immerso all'interno del campo operatorio e può effettuare queste attività a distanza dalla sala operatoria.

¹⁷ Cfr. un precedente parere del CNB e CNBBSV, *Sviluppi della robotica e della roboetica* (2017).

routine, o che li espongono a pericoli evitabili, permettendo loro di avere meno rischi e più disponibilità di tempo per il paziente.

L'automazione nell'acquisizione e interpretazione dei dati, nell'elaborazione delle diagnosi e nell'individuazione delle terapie o nella stessa effettuazione degli interventi non può essere completamente indipendente dall'uomo, ma richiede una costante verifica, pertanto non esclude la specificità della relazione tra medico e paziente. Non è possibile dimenticare che ogni soggetto è malato "a suo modo" e che il contatto personale costituisce l'elemento essenziale di ogni diagnosi e di ogni terapia. In questo senso, la macchina non potrà sostituire l'umano in una relazione che si costruisce sull'incontro di ambiti complementari di autonomia, competenza e responsabilità.

L'IA va considerata esclusivamente come un aiuto nelle decisioni del medico, che rimangono controllate e supervisionate dall'uomo. Resta compito del medico, in ogni caso prendere la decisione finale, in quanto la macchina fornisce solo ed esclusivamente un supporto di raccolta e analisi dei dati, di natura consultiva. Un sistema di "assistenza cognitiva automatizzata" nella attività diagnostica e terapeutica non è un "sistema decisionale autonomo". Esso effettua la raccolta di dati clinici e documentali, li confronta con statistiche relative a pazienti simili, accelerando il processo di analisi del medico.

Emerge un problema: cosa succede quando l'IA dimostra di avere una migliore performance rispetto al medico? In alcune circostanze questo è tecnicamente possibile ed è necessario tenere conto di questa evenienza. È in questo specifico spazio che potrebbe avvenire, nel futuro, la temuta 'sostituzione' della macchina all'uomo.

Ma un'ulteriore conseguenza, più immediata, può essere la delega delle decisioni alla tecnologia. Delegare compiti complessi a sistemi intelligenti, può portare alla perdita di qualità umane e professionali. Se la relazione di cura va configurata come relazione di fiducia, oltre che di cura (legge 219/2017), va preservato un ruolo sostanziale del "medico umano" in quanto unico a possedere le capacità di empatia e di vera comprensione, che non possono essere espresse dalla IA e che sole possono rendere reale un tale rapporto. Come suggeriscono alcuni dei maggiori studiosi di questi problemi bisogna garantire le quattro principali componenti: *Deep Phenotypizing*, *Deep Learning*, *Deep Empathy and Connection*.

La predeterminazione di canoni di comportamento e codici di condotta, come protocolli e linee guida, costituisce un supporto di conoscenze ed esperienza all'attività professionale, ma le esigenze di diagnosi e di cura spesso impongono di andare oltre i modelli prefissati. Sarebbe estremamente grave se lo spazio lasciato alla presunta neutralità delle macchine conducesse alla "neutralizzazione" del paziente. Le enormi potenzialità offerte dall'IA vanno considerate come la preziosa opportunità con cui la tecnica allarga gli orizzonti dell'etica, consentendo di accrescere gli spazi di ascolto del paziente e il contatto con il decorso della sua malattia. In questo senso l'IA sarebbe un utilissimo strumento che consente di risparmiare il tempo che il medico impiega in operazioni di routine in modo da ricavare più tempo per la relazione con il paziente.

3.2. L'affidabilità della IA e la opacità degli algoritmi

L'IA, come detto, è costituita da una serie di algoritmi: istruzioni precise ed espressioni matematiche per trovare associazioni, identificare tendenze, estrarre

dinamiche a partire dai dati raccolti e inseriti. Quando gli algoritmi operano, vengono considerati ‘affidabili e neutrali’ di per sé, solo per il fatto che i loro metodi sono rappresentati attraverso sistemi matematici, misurabili¹⁸.

Ma bisogna ricordare che è l’uomo (con l’aiuto della macchina) che raccoglie e seleziona i dati, e che costruisce algoritmi. In questo senso il sistema di IA può essere “opaco”. Per “opacità” si intende: che i passaggi attraverso cui si interpretano i dati non sempre sono spiegabili (trasparenti) e che possono anche dare risultati discriminatori. La discriminazione non deriverebbe dalla macchina ma dall’uomo che seleziona i dati ed elabora algoritmi. Questo passaggio implica una riflessione sull’“etica dei dati” a supporto dell’IA (che devono essere di qualità e interoperabilità) e sull’“etica degli algoritmi” (detta anche “algor-etica”), che dovrebbero essere basati su dati non selezionati oppure su selezioni inclusive e non discriminatorie.

L’IA, per quanto possa raggiungere un elevato grado di accuratezza, non è e non può essere sempre esplicativa. È impossibile da parte degli stessi programmatori e tecnici spiegare come il sistema sia arrivato a determinati risultati (black box problem). L’automazione può comportare la mancanza di trasparenza sulla logica seguita dalla macchina: la macchina non fornisce, né è possibile tracciare, informazioni sulle correlazioni indicate o sulla logica adottata per giungere a una conclusione o proporre una decisione (rivolta al medico e/o al paziente).

L’opacità sugli elementi essenziali e sul processo con cui un sistema di IA è arrivato ad una conclusione decisionale, comporta il rischio che l’operatore sanitario non possa convalidare e confermare, o ragionevolmente scartare, la proposta del sistema nel tentativo di prendere la propria decisione. Per un essere umano è praticamente impossibile analizzare l’enorme quantità di calcoli compiuti dall’algoritmo e scoprire esattamente come la macchina è riuscita a decidere. Questo solleva un problema per il medico nel confronto con la macchina (decidere se affidarsi o meno agli algoritmi) e nei confronti del paziente, al quale non può fornire una spiegazione e una informazione trasparente.

Gli algoritmi inoltre, con la classificazione di persone in gruppi o sottogruppi con profili simili a quelli associati a determinati schemi (clusterizzazione), può non tenere conto delle varianti che può presentare un particolare paziente. Una decisione di cura basata esclusivamente su profili elaborati sui pazienti e in modo automatizzato (algoritmi) può portare all’esclusione del trattamento senza offrirne in cambio un altro presunto meno efficace, ma indicato come alternativa. Emerge il rischio che, classificando o stratificando i pazienti in gruppi o sottogruppi in base ai profili personali da loro ottenuti sulla base di vari criteri o finalità, vengano assunte decisioni discriminatorie, stigmatizzanti o arbitrarie esclusivamente sulla base di tali profili o sulla base di considerazioni extra-sanitarie (legate anche indirettamente, ad esempio, all’origine etnica o al genere). Le “discriminazioni algoritmiche” sono possibili, anche in ambito medico, con un impatto su equità e inclusività. Le iniquità già esistono nell’ambito della salute, ma l’IA potrebbero accentuarle e peggiorarle creando e/o aumentando il “divario” e le disuguaglianze. È possibile evitare questa deriva con un approccio largo e rappresentativo dei dati utili, continuamente aggiornati, per lo sviluppo degli algoritmi.

¹⁸ Come già evidenziato non sempre è possibile ridurre l’IA a una logica sintetica come quella di un’equazione, proprio perché l’IA è più vicina a un concetto dinamico nel quale le singole osservazioni vengono paragonate agli atomi che costituiscono la base di conoscenza.

Inoltre, non si deve dimenticare che l'assistenza medica comporta anche grandi interessi economici¹⁹, pertanto l'IA può essere orientata, attraverso la costruzione degli algoritmi, ad influenzare in vari modi le decisioni del medico ad esempio facilitando le prescrizioni attraverso un aumento o una diminuzione dei valori di normalità per una serie di parametri funzionali o biochimici. Pertanto, l'IA può indurre a privilegiare una classe di farmaci rispetto ad una altra classe che abbia le stesse indicazioni per un determinato sintomo o patologia. Può privilegiare un percorso diagnostico che favorisca l'utilizzo di determinati reattivi piuttosto di altri. Può suggerire l'impiego di determinate apparecchiature e tecnologie più costose in alternativa ad altre più economiche. Può influenzare il medico a prescrivere trattamenti piuttosto che a stimolare il paziente a migliorare buoni stili di vita²⁰.

Queste fra le altre ragioni di rischio spingono il Comitato a ritenere che debbono essere fatti controlli accurati, anche attraverso la validazione degli algoritmi, in modo da ottenere la più probabile certezza che l'introduzione delle varie forme di IA siano vantaggiose per migliorare la qualità delle prestazioni del Servizio Sanitario Nazionale. In altre parole tutti i "prodotti" della IA devono venire confrontati, attraverso studi condotti con le regole degli studi clinici controllati, con le decisioni che vengono assunte indipendentemente dalla IA da parte di gruppi di medici competenti²¹. Fermo restando che gli studi clinici controllati rimangono il "gold standard" per la dimostrazione dell'efficacia e della sicurezza dei trattamenti, bisogna tenere conto che quando si parla di applicazione di IA in medicina si fa riferimento a un software²². Con il problema che il meccanismo cambia nel tempo e la validazione richiede un monitoraggio e ulteriori controlli.

Solo se emerge da questi studi che la IA ha una performance migliore di quella dei medici, dovrebbe essere accettata ed utilizzata. Bisogna però considerare il significato di "performance": ad esempio, la segmentazione di immagini diagnostiche può essere fatta con elevata qualità da un medico, ma un sistema di IA impiega una frazione di tempo per fare la stessa operazione e non si stanca mai.

Ciò ha particolare rilievo per migliorare la qualità delle prestazioni del Servizio Sanitario Nazionale nell'interesse del cittadino. È pure essenziale che siano istituite anche strutture, adatte alla ricerca con una mentalità volta all'interesse pubblico, a farsi carico dello sviluppo della IA attraverso fondi pubblici. Bisognerà dunque dimostrare la sicurezza della IA sulla base del controllo che è la base dei dati da cui si parte, il vantaggio in termini di benefici e rischi, in senso clinico, e costo-efficacia, nonché la diffusibilità e la sostenibilità di tali tecnologie sul territorio e nel tempo. Solo in questo modo si potrà dimostrare l'affidabilità di tali sistemi mediante certificazioni che ne garantiscano l'utilizzabilità nella pratica clinica. Solo in questo modo si potranno affidare

¹⁹ Basti ricordare che il solo mercato dei farmaci vale almeno 30 miliardi di euro a cui si aggiunge il mercato degli strumenti diagnostici, dei dispositivi medici e riabilitativi.

²⁰ Cfr. R. SPARROW, J. HATHERLEY, *High Hopes for "Deep Medicine"? AI, Economics, and the Future of Care*, "The Hastings Center Report", 18 febbraio 2020.

²¹ Ad esempio, se una *learning machine* è programmata per fare diagnosi e terapia per malattie polmonari deve essere valutata rispetto alle decisioni di un gruppo di medici con specializzazioni in pneumologia.

²² È lo stesso tema della *digital therapeutics* che sta spingendo verso alcune riflessioni (*real world evidence*) che hanno anche implicazioni regolatorie. Oltre la questione metodologica esiste ancora incertezza su quale ente regolatorio dovrà trattare il tema. È in corso una discussione sia ad EMA ma anche ad AIFA.

mansioni anche complesse in maniera da supportare il rapporto fiduciario tra paziente e IA.

3.3. IA e dati: tra privacy e condivisione dei dati

IA in medicina “si alimenta” di dati: i dati sono indispensabili per il “training” della macchina e sono gli elementi base della costruzione di algoritmi, modelli matematici che interpretano i dati. La disponibilità di dati (dati clinici, immagini, dati genetici, ecc.), la accuratezza e qualità dei dati, la interoperabilità dei dati (mediante criteri di standardizzazione e classificazione) sono le condizioni necessarie per gli sviluppi e applicazioni della IA. Poiché ogni sistema di IA si basa sui dati emerge il problema della verifica, selezione, preparazione e supervisione dei dati da esseri umani, evitando gli errori della raccolta e classificazione dei dati, prevedendo anche meccanismi di IA di controllo e verifica della correttezza.

L’insistenza sulla protezione della privacy e confidenzialità è spesso sottolineato come un ostacolo allo sviluppo della IA. Coloro che intendono applicare la IA insistono sulla necessità di disporre di dati in un vasto ambito di azione, sul piano globale (dunque anche trasferimento di dati in altre nazioni) e conservazione dei dati, ma anche di campioni e immagini associate, nel tempo. Dati non pienamente anonimi ma pseudonomizzati che consentano una tracciabilità, identificazione in casi di rilevanza di comunicazione dei risultati, con opportune condizioni per evitare rivelazioni inappropriate.

La enorme raccolta di dati, necessari alla IA, fa emergere anche il rischio legato nell’uso dei dati e nell’incrocio dei dati, alla re-identificazione sia intenzionale che accidentale, sollevando il problema della privacy, che in tale contesto tende a “evaporare”²³. Al punto che si ritiene che le tecnologie tendono a divenire sempre più “opache” e gli utenti “trasparenti”.

Nell’epoca della IA e della esigenza dell’uso di dati per la ricerca in medicina, ci si interroga sulla possibilità della “condivisione” dei dati (*data sharing*) come “bene sociale” per l’avanzamento delle conoscenze scientifiche.

Esistono metodi e tecnologie per eseguire le transazioni di dati preservando la sicurezza dei dati (una delle tecnologie è la famiglia delle applicazioni *block-chain*).

Una condivisione che comunque sia garantita da un uso esclusivo per fini di ricerca, che consenta un ritorno di informazione e una condivisione dei risultati clinicamente rilevanti (*benefit sharing*).

Ampio è il dibattito, anche sul piano regolatorio, della applicabilità della GDPR agli scenari della IA, ove risulta essere irrealistico proteggere la privacy e garantire il controllo dei dati, nello spazio globale della ricerca (ICT) e nei tempi non definibili a priori.

²³ Sul tema della privacy il CNB è intervenuto nel documento “*Tecnologie dell’informazione e della comunicazione e big data: profili bioetici*” (2016) sottolineando come nell’ambito del “trattamento dei dati nel momento della richiesta delle informazioni, che deve essere sempre accompagnata da un consenso informato esplicito”, in modo trasparente, completo e semplice, specificando “chi raccoglie e chi userà i dati, quali dati, come vengono raccolti, dove verranno conservati e per quanto tempo, per quale ragione e per quale scopo”, precisando la revocabilità. Nel parere “*Mobile-health e applicazioni per la salute: aspetti bioetici*” (2015), il CNB esprime la consapevolezza “della difficoltà di realizzare un consenso informato e di tutelare la privacy degli utenti in questo nuovo campo applicativo”.

3.4. Consenso e autonomia

Il rapporto favorevole efficacia/rischi e il requisito del consenso informato e dell'autonomia sono regole fondamentali nel rapporto paziente-medico/operatore sanitario, in quanto tutelano il diritto alla vita, alla salute, alla dignità della persona, all'autodeterminazione.

Ne consegue che questo debba essere centrale anche nel caso che il medico voglia avvalersi nel trattamento di cura di dati ricavati dall'IA e dello sviluppo della robotica. Ma l'informativa è tutt'altro che facile da dare e l'autonomia/consenso si complica con l'IA che suscita un senso di disorientamento in considerazione della velocità con cui le tecnologie stanno stravolgendo il mondo conosciuto. Non è facile per la persona-paziente immaginare quali conseguenze potrebbero scaturire da questi nuovi progressi tecnologici: È il medico che deve mediare questa comunicazione. Terminologie complesse, parole che possono suonare misteriose, si riscontrano nelle nuove procedure sanitarie (*machine learning, deep learning, neural networks, big data, algoritmi, cloud, ecc.*) che rendono il consenso ai nuovi trattamenti sanitari sempre più complesso e forse dato più sulla fiducia verso il medico che non sulla effettiva comprensione. Il consenso informato a trattamenti sanitari basati su IA può influire sull'autonomia del paziente. Certamente il paziente vede modificarsi il tradizionale rapporto di alleanza con il medico: ha ancora idee molto confuse sulle applicazioni dell'IA, ne verifica i vantaggi, ma non ne comprende appieno i rischi.

È, dunque, un obbligo etico e giuridico che coloro che si sottopongono a trattamenti sanitari così innovativi, attraverso l'IA, siano informati nelle modalità più consone e comprensibili al paziente di ciò che sta accadendo, di essere (se è il caso) oggetto di sperimentazione e validazione; di essere a conoscenza che ciò che è loro applicato (sul piano diagnostico e terapeutico) implica dei vantaggi, ma anche dei rischi. Va specificato in modo esplicito nel consenso informato se i trattamenti applicati (diagnostici o terapeutici) provengano solo da una macchina (IA, robot) o se e quali sono gli ambiti e i limiti del controllo umano o supervisione sulla macchina. Queste difficoltà di una comprensibile ed esauriente informativa data dal medico al paziente (difficoltà sia di comunicazione del medico che di ricezione del paziente) nel momento in cui ci si avvale di trattamenti che fanno uso dell'IA è accresciuta dalla opacità degli algoritmi.

3.5. La responsabilità

L'automazione in medicina può contribuire alla riduzione degli incidenti e della mortalità (aumentare attenzione e accuratezza delle azioni del medico, potenziare il suo impiego anche nelle procedure di routine, ecc.), ma, come detto in precedenza, non è priva di rischi. Le macchine possono essere mal programmate e mal impiegate. Pertanto, la questione della responsabilità è uno dei problemi più delicati e complessi che sorgono con l'utilizzo e lo sviluppo dei nuovi sistemi di IA. In particolare il problema riceve una sempre maggiore attenzione a livello di policy e di strategia legislativa. Il tentativo è soprattutto quello di fare chiarezza se la responsabilità per alcune decisioni prese attraverso un sistema intelligente debbano essere attribuite al progettista, al venditore del software, al proprietario, all'utilizzatore (il medico) o a terze parti. L'eventuale evenienza di incidenti dovrebbe essere tracciata e analizzata come avviene per qualsiasi errore medico.

Ogni evoluzione che nel campo medico finisca per modificare il rapporto paziente-medico, dovrebbe prevedere un intervento del diritto che governi le innovazioni in coerenza con il "sistema" esistente, realizzando così una garanzia sia per il paziente, quanto a copertura dei nuovi rischi, sia al contempo per l'operato del medico. Il fattore che pone nuove istanze di mediazione giuridica non è tanto la presenza di una intelligenza con eventuale auto-apprendimento, quanto la circostanza che l'IA ha un "autore" che la crea e che può non coincidere con il "produttore" del bene che la incorpora, il "venditore" e l'utilizzatore e per i quali si pone il problema di delineare diritti, limiti e responsabilità. In molteplici casi il comando dell'azione rimane al medico, ma questo non è direttamente l'agente dell'azione, anzi può occupare a volte una posizione distante nello spazio. Il fatto che vi sia una "catena di comando" cui far risalire la responsabilità dell'azione può far pensare che l'azione sia meno soggetta al caso e all'improvvisazione, ma ogni anello della catena ha il suo punto di fragilità e data la complessità della struttura gestazionale e dell'azione non è scontato dire chi alla fine ha la responsabilità di quanto eventualmente accade a danno del paziente o se questa responsabilità è unica. Queste autonome e distinte responsabilità dovrebbero poter essere fatte direttamente valere dal fruitore finale di quel prodotto, il paziente, non solo attraverso il tradizionale piano contrattuale che vincola il medico e la struttura sanitaria. A tali responsabilità si aggiungono quelle distinte nell'ambito dei rapporti tra professionisti (progettista, validatore, venditore del software, programmatore, ecc.), che hanno contribuito al formarsi della catena paziente-medico, senza che vi sia stata una preventiva proposta, un preventivo atto formale, cosicché è davvero difficile ravvisare egualmente un contratto. Siamo in una logica che ci avvicina, alla categoria, elaborata dalla giurisprudenza, della "responsabilità da contatto sociale", che ipotizza una obbligazione che si collega al dovere di diligenza nell'osservare le regole dell'arte che si professa.

Sono aspetti tutti questi nei quali l'analisi delle nuove responsabilità e delle ipotizzabili nuove norme, così come delle nuove interpretazioni evolutive di norme esistenti, rendono indispensabile la collaborazione fra scienze giuridiche e scienze mediche, dovendo le prime confrontarsi con le seconde e viceversa. È opportuno ed anzi necessaria una riflessione continua interdisciplinare che veda le due competenze "parlarsi fra loro", anche ai fini di delineare il futuro assetto delle possibili molteplici responsabilità mediche connesse con la IA.

3.6. La formazione in ambito medico, tecnologico e sociale

Il mondo medico e degli operatori sanitari non è oggi pienamente, salvo eccezioni, addestrato ad utilizzare in modo responsabile i risultati della ricerca di IA. È perciò molto importante agire su due fronti: da un lato inserire le problematiche derivanti dall'IA nell'ambito delle attività della Educazione Continua Medica (ECM) svolta in modo indipendente e dall'altra una riforma delle Scuole di Medicina, oltre che nelle scuole degli operatori sanitari.

L'inserimento dell'IA nella educazione dei medici e degli operatori sanitari²⁴ rientra nel c.d. *reskilling* dei dipendenti, ossia riconvertire i lavoratori (in questo caso gli operatori sanitari) di fronte agli sviluppi delle tecnologie emergenti.

²⁴ In questi anni l'Europa sembra avere preso sempre più coscienza dell'importanza del problema, così come cominciano ad interessarsene diversi comitati che stanno delineando gli orientamenti della riflessione etica e giuridica nel campo della IA.

Tuttavia, formare il personale sanitario dipendente per occupare le stesse posizioni, ma che implicano nuove necessità e professionalità, sarà più complicato e dispendioso che creare nuovi posti di lavoro a favore di persone già formate alla comprensione della IA nel campo medico²⁵. Ne consegue la preoccupazione dell'European Group on Ethics in Science and New Technologies (EGE), nel parere *Future of Work, Future of Society 2018* della "polarizzazione delle abilità" (*skills polarization*) che può nascondere nuove forme di discriminazione, escludendo coloro che non sono in grado di apprendere le nuove "abilità" richieste. Il problema delle nuove professioni, anche nel campo medico, resta, dunque, quello che si richiederanno competenze di livello elevato. Questa discussione rientra nel campo delle nuove diagnosi e terapie; l'aggiornamento continuo è essenziale per i medici e operatori sanitari.

L'altro percorso è la reimpostazione dei programmi di educazione medica, destinando una parte significativa della formazione del futuro medico alle problematiche derivanti dalla digitalizzazione della medicina che sta alla base delle tecnologie di IA di cui il futuro medico dovrà usufruire essendo in grado di comprenderne vantaggi, limiti e pericoli. È auspicabile la istituzionalizzazione di corsi interdisciplinari per la formazione dei professionisti della salute ad un adattamento costante al cambiamento tecnologico e alla possibile "convergenza" e trasversalità dei settori disciplinari tradizionali (es. medicina e informatica o fisica o data science con fondamenti di computer science e IA, componenti di etica clinica, bioetica e biodiritto).

La formazione dovrà essere rinnovata anche nell'ambito della tecnologia, introducendo corsi di formazione all'etica e alla bioetica per gli ingegneri, informatici, computer scientist, con particolare riferimento all'etica nel disegno delle tecnologie (*ethics by design/in design/for designers*) e nella progettazione, metodologia e applicazione delle tecnologie. Solo in questo modo sarà possibile garantire una consapevolezza etica di coloro che costruiscono le tecnologie, per consentire che i principi e valori siano presenti sin dall'inizio della progettazione tecnologica.

È auspicabile anche la promozione di un dibattito pubblico sugli sviluppi e limiti della IA in medicina, affinché sia possibile da parte di tutti i cittadini la acquisizione degli elementi di "AI literacy", la partecipazione attiva alla discussione sociale. Sul lungo termine è auspicabile che nella scuola la introduzione della scienza, come parte essenziale della cultura, possa porre le basi per una comprensione della presenza della IA nei diversi settori. Sono queste le premesse indispensabili per un possibile superamento del "divario digitale" in medicina, promuovendo una maggiore inclusività.

4. Raccomandazioni

Il CNB intende richiamare, alla luce della precedente analisi, alcuni principi etici di riferimento nell'ambito dell'uso della IA in medicina²⁶. A fronte di un progresso in questo ambito considerato "trasformativo" e "dirompente", soprattutto nell'ambito della tutela della salute, il Comitato intende promuovere una riflessione etica nel bilanciamento della dimensione umana e della dimensione artificiale, senza la reciproca esclusione. Nella convinzione che:

²⁵ Da ciò consegue che nonostante l'emergere di molte nuove professioni, potremmo essere testimoni del formarsi, come già da più parti ipotizzato, di una classe di individui disoccupati e "inutili".

l'esclusione dell'artificiale toglie molte opportunità all'uomo; l'esclusione dell'umano solleva molte criticità dati i limiti dell'artificiale. Bisogna evitare eccessive speranze, ma anche eccessivi timori, con un atteggiamento di fiducia e di cautela.

Il Comitato raccomanda:

- di predisporre *ex ante* accurati controlli per l'"addestramento" delle macchine sulla base di dati di qualità, aggiornati e interoperabili e di condurre sperimentazioni adeguate nell'ambito della IA per garantire sicurezza ed efficacia nell'uso di queste nuove tecnologie e sollecitare la ricerca di strumenti di validazione e certificazione delle tecnologie e di sorveglianza e monitoraggio, come elementi indispensabili per creare un "patto sociale di fiducia e affidabilità" delle tecnologie in ambito medico; sarebbe auspicabile integrare nell'ambito dei comitati etici per la sperimentazione la figura di un computer scientist o un esperto di IA, e aggiornare anche la normativa sulla sperimentazione con riferimento a software in ambito medico;

- nell'ambito della relazione medico-paziente, informare, soprattutto in questo periodo di transizione, in modo corretto i malati dei rischi e benefici dell'uso della IA con riferimento alle specifiche applicazioni (e anche dei limiti di spiegabilità delle tecnologie "opache"), al fine di garantire la piena consapevolezza delle scelte e assicurando anche percorsi alternativi nella misura in cui emergesse una resistenza all'accettazione delle nuove tecnologie; garantire, nelle applicazioni della IA alla salute, un approccio largo e rappresentativo (non selettivo e discriminante) e un ambito di "controllo umano significativo" di una interazione e collaborazione uomo/macchina, a tutela della correttezza complessiva e della comunicazione paziente-medico come ambito della cura;

- ripensare la formazione dei professionisti della salute in modo dinamico con una revisione flessibile dei programmi di studio da parte di commissioni interdisciplinari, per un adattamento costante al cambiamento tecnologico, pensando anche a percorsi di possibile "convergenza" dei settori disciplinari tradizionali (es. nella facoltà di medicina, medicina e informatica o fisica o data science e simmetricamente, nella facoltà di giurisprudenza/scienze umane con fondamenti di computer science e IA);

- introdurre la rilevanza dei principi etici di autonomia, responsabilità, trasparenza, giustizia nei codici di condotta e nei corsi di formazione degli ingegneri, informatici, sviluppatori, con particolare riferimento all'etica nel disegno delle tecnologie (*ethics by design/in design/for designers*), assicurando una tecnologia che sia orientata ad incorporare i valori e assicurare la centralità del paziente;

- creare una coscienza pubblica della società su opportunità e rischi delle nuove tecnologie, affinché i cittadini possano partecipare in modo critico al dibattito su IA senza un affidamento cieco e nemmeno un eccesso di preoccupazione, essere consapevoli delle scelte e delle implicazioni della sanità digitale: tale promozione può avvenire anche mediante l'organizzazione di conferenze per le scuole e incontri con la cittadinanza, che il CNB regolarmente propone;

- sollecitare, sul piano normativo, un aggiornamento sui profili concernenti la responsabilità nell'applicazione delle nuove tecnologie;

- promuovere la ricerca su IA, oltre al settore privato, anche e soprattutto nella dimensione pubblica del Servizio Sanitario Nazionale (SSN).